

A comparison of geographical information systems–based algorithms for computing the TOPMODEL topographic index

Feifei Pan

Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

Christa D. Peters-Lidard

Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

Michael J. Sale and Anthony W. King

Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

Received 2 February 2004; revised 1 March 2004; accepted 20 April 2004; published 24 June 2004.

[1] The performance of six geographical information systems (GIS)-based topographic index algorithms is evaluated by computing root-mean-square errors of the computed and the theoretical topographic indices of three idealized hillslopes: planar, convergent, and divergent. In addition to these three idealized cases, two divergent hillslopes with varying slopes, i.e., concave (slopes decrease from top to bottom) and convex (slopes increase from top to bottom) are also tested. The six GIS-based topographic index algorithms are combinations of flow direction and slope algorithms: i.e., single flow direction (SFD), biflow direction (BFD), and multiple flow direction (MFD) plus methods that determine slope values in flat areas, e.g., W-M method [Wolock and McCabe, 1995] and tracking flow direction (TFD) method. Two combinations of horizontal resolution and vertical resolution of the idealized terrain data are used to evaluate those methods. Among those algorithms the MFD algorithm is the most accurate followed by the BFD algorithm and the SFD algorithm. As the vertical resolution increases, the errors in the computed topographic index for all algorithms decrease. We found that the orientation of the contour lines of planar hillslopes significantly influences the SFD's computed topographic index. If the contour lines are not parallel to one of eight possible flow directions, the errors in the SFD's computed topographic index are significant. If mean slope is small, TFD becomes more accurate because slope values in flat areas are better estimated. **INDEX TERMS:** 1899 Hydrology: General or miscellaneous; 1824 Hydrology: Geomorphology (1625); 1832 Hydrology: Groundwater transport; **KEYWORDS:** GIS, TOPMODEL, topographic index, single flow direction algorithm, biflow direction algorithm, multiple flow direction algorithm

Citation: Pan, F., C. D. Peters-Lidard, M. J. Sale, and A. W. King (2004), A comparison of geographical information systems–based algorithms for computing the TOPMODEL topographic index, *Water Resour. Res.*, 40, W06303, doi:10.1029/2004WR003069.

1. Introduction

[2] TOPMODEL is a topography-based concept for watershed hydrology modeling. Since the TOPMODEL was first proposed in 1979 [Beven and Kirkby, 1979], it has been widely used to study the effects of topography on hydraulic processes including flood frequency, streamflow generation, flow paths, geomorphic characteristics, and water quality [Wolock and McCabe, 1995]. In addition to the success of the TOPMODEL concept in traditional hydrologic modeling, this concept has also been successfully incorporated into several ecosystem-atmosphere models including the Regional Hydro-Ecological Simulation System (RHESSys) [e.g., Ford *et al.*, 1994]; the Land Ecosystem-Atmosphere Feedback model (LEAF-2) [Walko *et al.*, 2000]; the TOPMODEL-based Land Atmosphere Transfer Scheme

(TOPLATS) [Famiglietti and Wood, 1994; Peters-Lidard *et al.*, 1997]; the Catchment Model [Koster *et al.*, 2000]; and the Common Land Model (CLM) [Dai *et al.*, 2003].

[3] To apply the TOPMODEL, a modeled catchment is partitioned with a regular grid or lattice. The so-called 'topographic index' is then calculated for each cell in the catchment. The topographic index, $\ln(a/\tan \beta)$, is the natural logarithm of the ratio of the specific flow accumulation area a to the ground surface slope $\tan \beta$. The surface slope can be evaluated from digital elevation model (DEM) data. The specific flow accumulation area is the total flow accumulation area (or upslope area) A through a unit contour length L . To compute the total flow accumulation area A , flow directions are tracked upslope, starting from the cell of interest to the upstream divide of the watershed, and then tracked downslope accumulating cells contributing to the drainage area of the cell of interest. Here we note an uncertainty associated with the definition of the flow accumulation area A . From the presentations of Beven and

Wood [1983, Figure 2] and Kirkby [1997, Figure 1], one can reasonably conclude that the flow accumulation area is defined along the ground surface. However, using DEMs in geographical information systems (GIS), the computed flow accumulation area is generally the area projected to x-y plane, and this calculation of A has become standard practice. The difference between these two areas is negligible if the slope is less than 0.5 (m/m), and most of the slopes in the watersheds to which the TOPMODEL is applied are less than 0.5 (m/m) [e.g., Montgomery and Dietrich, 1992]. Therefore, for consistency with standard practice, we adopt the convention of calculating the flow accumulation area or upslope area as the area projected to x-y plane.

[4] The slope term ($\tan \beta$) in the topographic index arises from the assumption that the surface of the water table is parallel to the ground surface. Thus the local hydraulic gradient is assumed to be equal to the slope of the ground. Because flow direction depends on hydraulic gradient, or ground surface slope, flow direction and the calculation of the upslope accumulation area should be consistent with the local slope value that is used to compute $\ln(a/\tan \beta)$. Thus the computed topographic index is dependent upon the calculation of both slope and flow direction.

[5] Although many researchers have investigated algorithms for calculation of slope and flow direction, those studies tend to focus on either slope [e.g., Jones, 1998; Zhang et al., 1999] or flow accumulation area [e.g., Tarboton, 1997; Rieger, 1998]. Few studies have examined the combined effects of slope and flow direction algorithms on the topographic index. The works of Quinn et al. [1991], Wolock and McCabe [1995], and Mendicino and Sole [1997] are exceptions. However, these studies only investigated the difference in the statistical moments or distributions of the computed topographic index; no comparisons of errors between the “true,” i.e., analytically solved, and the numerically computed topographic indices were carried out. Comparing against “truth” is difficult because an analytical expression for the real terrain does not normally exist. Therefore one cannot usually determine which topographic index algorithm is more accurate. This paper seeks to resolve that issue and provide objective evaluation of the appropriate numerical algorithm for calculating the topographic index.

2. Topographic Index Algorithms

[6] A topographic index algorithm is actually a combination of two algorithms, one to calculate flow direction and another to calculate slope. Here we describe three commonly used flow direction algorithms that allow flow in one, two, and more than two directions. We identify each topographic index algorithm by the name of the associated flow direction algorithm (e.g., single flow direction, biflow direction, and multiple flow direction), because each flow direction algorithm uses its own algorithm to determine slope.

2.1. Single Flow Direction (SFD) Algorithm

[7] The single flow direction (SFD) algorithm [O’Callaghan and Mark, 1984] calculates flow direction as the steepest slope direction, which is determined by the Maximum Downward Gradient (MDG). This SFD algorithm, also known as the D8 algorithm, is widely used in DEM data analysis [e.g., O’Callaghan and Mark, 1984; Band, 1986; Greenlee, 1987; Mark, 1988; Jenson and Domingue,

1988] and GIS software (e.g., the “FLOWDIRECTION” function in ARC/INFO GRID).

[8] MDG computes the downhill elevation gradients of a 3×3 cell window along eight directions (i.e., elevation of the center cell minus elevation of each of its eight neighbors divided by the distance between those two cells). The slope of the central cell is calculated as the largest of the eight directions. Steepest slope direction is the direction from the central cell to the neighbor generating the largest downhill elevation gradient.

[9] If the central cell has a lower elevation than one of its neighbors, the downhill gradient along this direction is negative. If the calculated slope (i.e., the largest one among eight directions) is less than zero, this cell is called a sink or a pit. At a sink, there is only inflow, and no outflow. However, it is often computationally required to force watersheds to have outlets. Therefore sinks in DEM data are usually filled before watersheds can be delineated and other hydrologic parameters estimated. One common filling approach for sinks generated by MDG is to raise the elevation of a sink to the lowest elevation among its neighbors [e.g., Jenson and Domingue, 1988; Tarboton et al., 1991]. This algorithm is used in the ARC/INFO GRID (i.e., the “FILL” function). For example, the recalculated slope value of the “filled” cell is now zero; the cell and its neighbors form a flat area. To determine flow direction in a flat area, the method suggested by Jenson and Domingue [1988] (e.g., the ARC/INFO FLOWDIRECTION function) is used.

[10] Once the single flow direction is determined for each cell, flow accumulation area (or upslope area) (A) is calculated using a recursive procedure [e.g., Tarboton, 1997]. The specific flow accumulation area (a) is (A) divided by a contour length, which is equal to the grid size or horizontal resolution of the DEM. The slope is set to be the maximum downward elevation gradient.

6. Conclusions

[39] By comparing the numerically computed and theoretical (analytically calculated) topographic indices for 526 cases, we found that the MFD algorithm is best in terms of accuracy for all idealized hillslope cases. Although the BFD algorithm is a little worse than the SFD algorithm when the angle between the contour line and x axis is 45° for planar hillslopes, overall the BFD algorithm is better than the SFD algorithm. The difference in the errors of the computed topographic index due to the different algorithms for assigning the contour lengths is negligible (i.e., MFD [Quinn et al., 1991] and MFD* [Wolock and McCabe, 1995]). We found that the orientation of the contour lines of the planar hillslopes significantly influences the SFD’s computed topographic index. If the contour lines are not parallel to one of eight possible flow directions, the errors in the SFD’s computed topographic index are significant.

[40] Recalculating nonzero slopes in flat areas increases errors only when the mean slope value is less than 0.5VR/HR. The tracking flow direction (TFD) method is more accurate than the W-M method [Wolock and McCabe, 1995] because TFD can give a minimum slope value less than 0.5VR/HR. Therefore, to achieve the highest accuracy, we recommend using a combination of the MFD algorithm and the TFD method to compute topographic index, especially when the vertical resolution of DEM data is low.